MANIKANDAN MUTHIAH

Mail id: pcbzmani@gmail.com   Mobile: +91 9500327540/+97451173177

PROFESSIONAL OBJECTIVE

Senior Data Engineer with 9+ years of experience designing, building, and supporting large-scale data pipelines across Azure and GCP. Strong expertise in Spark, PySpark, SQL, and cloud-native data platforms, with proven airline domain experience and a strong ownership-driven engineering mindset.

TECHNICAL SUMMARY

- Around 9 years of Software industry experience. Extensively in Development projects with multiple deployments, Go-Live implementation and upgrade.
- Around 8 years of experience in data engineering.
- Around 7 years of experience in Agile workflow model.
- Around 8 years of experience is Azure cloud services like ADF, Azure Databricks.
- Around 2 years of experience in GCP cloud services like Dataproc, Composer and Bigquery.
- Have extensive knowledge on Healthcare, Mortgage Insurance, Retail and Airline domain.

EDUCATION

- Bachelors of engineering in Electrical and electronics, National engineering college with 91%. Year 2012-2016.

CERTIFICATION

- Certified in Azure Data Engineer (DP 203)
- Certified in Databricks – Data engineer associate and Data analyst.
- Certified in Google Cloud Associate Data engineer.

JOB PROFILE

- Senior Software engineer at EMAKINA (EPAM Systems), Doha – Sep 2023 – Till date
- Senior Software engineer at EPAM Systems, Chennai – Apr2022 – Sep 2023
- Application development senior analyst at Accenture, Chennai – Apr2021 –April 2022.
- IT Analyst at TCS, Chennai – Jul 2016 – Apr 2021.

LinkedIn Profile: https://www.linkedin.com/in/manikandan-muthiah-6a096214b/

Portfolio   : https://pcbzmani.github.io/portfolio/

TECHNICAL SKILLS

Programming Languages	Python, SQL
Cloud Platforms	Azure, GCP
Azure & GCP Tools	Azure Data factory, Azure Databricks, Dataproc, Composer (Airflow), Bigquery
Database	MS SQL, Oracle, SQL Server, Postgressql, Delta tables

ROLE & RESPONSIBILITIES

- Maintained metadata, data dictionaries, and source-to-target mappings to ensure clarity, traceability, and alignment with enterprise data governance standards.
- Updated data inventories and ensured data lineage documentation for analytics datasets across Azure and GCP platforms.
- Implemented access controls and followed data security and privacy guidelines while designing and deploying data pipelines.
- Operated with a "You Code It, You Own It" mindset by supporting pipelines in production, handling incidents, and ensuring SLA adherence.

PROJECT DETAILS

Project: Duty Free Business Intelligence
Client: Leading Aviation in Qatar
Project Description:

A global aviation duty-free service company that offers tax-exempt shopping for luxury goods, travel essentials, and exclusive travel retail experiences, along with loyalty programs, primarily to international travellers at airports and key transit hubs worldwide.
Role: Senior Data Engineer
Environment: Azure Databricks, Google Cloud Platform (GCP), GCP BigQuery, Google Composer (Airflow), PowerBI
Duration: Sep 2023 – Present
Role & Responsibilities:

- Applied analytical data modelling techniques to design curated datasets optimized for BI and reporting use cases.
- Designed fact and dimension-style datasets to support sales, inventory, and operational analytics.
- Collaborated with analysts and BI teams to align data models with business KPIs and reporting requirements.

Project: Aviation Inventory Analytics
Client: Leading Aviation in Qatar
Project Description:

A data and analytics initiative focused on aviation inventory management, providing insights into booking predictions, transaction patterns, and comprehensive flight operations (Flight 360) to enhance operational efficiency.
Role: Senior Data Engineer
Environment: Azure Databricks, Delta Tables, Dremio
Duration: Jul 2022 – Sep 2023
Role & Responsibilities:

- Designed and developed an ingestion framework to load data from on-premise systems to Azure cloud, ensuring high availability and scalability.
- Enhanced performance of long-running jobs by applying best practices, such as Delta table optimizations and query tuning.
- Conducted peer code reviews to ensure adherence to coding standards, performance best practices, and maintainability.
- Mentored junior engineers and contributed to reusable ingestion frameworks and engineering best practices.

Project: Data Factory for CTC
Client: Retail & Distribution
Project Description:

A data warehousing solution to centralize and process retail and distribution data, enabling efficient inventory management and operational insights through cloud-based analytics.
Role: Senior Data Engineer

Environment: Azure Data Factory, Azure Databricks, Delta Lake, Azure MySQL, MSSQL, Bitbucket, Jenkins, qTest, JIRA, Confluence, HDFS
Duration: Apr 2022 – Jun 2022
Role & Responsibilities:
- Developed a file- and RDBMS-based ingestion framework to transfer data from on-premise systems to Azure cloud with minimal latency.
- Created and optimized pipelines in Azure Data Factory to extract data from on-premise sources and load it into Azure Databricks Delta Lake.
- Applied Spark optimizations, including broadcast joins, repartitioning, coalesce, and salting, to enhance processing efficiency.
- Implemented data quality checks, validations, and failure-handling mechanisms within ingestion and transformation pipelines.
- Designed pipelines with monitoring and alerting to proactively detect data delays, schema changes, and job failures.

Project: Home Buyer Transformation
Client: NA Mortgage Insurance
Project Description:
A program leveraging big data and AI to assist homebuyers with mortgage insurance, enabling rapid and informed business decisions through advanced analytics.
Role: Data Engineer
Environment: Azure Data Factory, Azure Databricks, Azure SQL, PySpark, Hive, Azure DevOps, VS Code
Duration: Apr 2021 – Apr 2022
Role & Responsibilities:
- Delivered code iteratively across sprints, ensuring timely feature deployment.
- Utilized Azure Data Factory (ADF) to ingest data from diverse sources into Azure Data Lake for centralized storage.
- Designed and implemented data pipelines in ADF, integrating Databricks for transformation and loading data into target databases.

Project: Healthcare Data Modernization
Client: US Healthcare
Project Description:
A global health insurance company offering health, dental, vision, life, and accidental insurance, along with Medicare and Medicaid products, modernizing legacy systems for enhanced data processing and reporting.
Role: Developer
Environment: MS SQL, IBM DB2, Sqoop, Hive, PySpark, HDFS, AWS Redshift, Python, Spark
Duration: Jul 2016 – Apr 2021
Role & Responsibilities:
- Converted legacy COBOL SQL code into optimized Hive queries, improving query performance.
- Created Hive tables to view and analyse data stored in HDFS, incorporating optimization techniques like bucketing and indexing.
- Implemented PySpark with DataFrames API and Spark SQL for faster processing, applying optimization techniques such as broadcast joins and repartitioning.

I do hereby declare that the information furnished above is true to the best of my knowledge and belief.

Yours faithfully,
Place: Doha, Qatar   Manikandan Muthiah